

3 - VARIABLES ALEATORIAS

3.1- Generalidades

En muchas situaciones experimentales se quiere asignar un número real a cada uno de los elementos del espacio muestral. Al describir el espacio muestral de un experimento un resultado individual no tiene que ser necesariamente un número, por ejemplo, al tirar una moneda y tomar como espacio muestral $S = \{c, s\}$, o al tirar un dado dos veces tomamos como espacio muestral a $S = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}$, aquí S es un conjunto de pares ordenados.

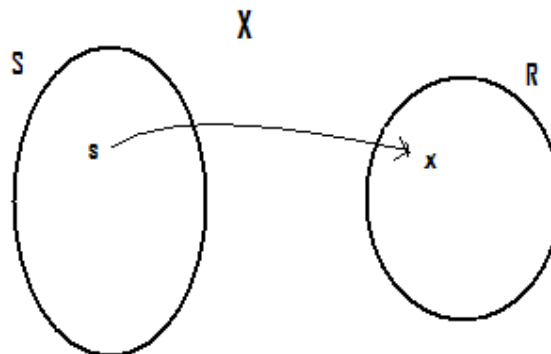
Definición: Sea \mathcal{E} un experimento aleatorio y S un espacio muestral asociado a él. Una **variable aleatoria** es una función que asigna a cada elemento de S un número real.

Notación: se anota a una variable aleatoria con letras mayúsculas X, Y, Z, W, \dots

Entonces, si X es una variable aleatoria de S en R

$$X : S \rightarrow R \quad \text{tal que} \quad X(s) = x$$

Con diagramas de Venn



Desde ahora en lugar de escribir variable aleatoria, escribiremos *v.a.*

Ejemplos:

1- Se tira una moneda tres veces

Sea X la v.a. X : “número de caras obtenidas luego de los tres tiros”

Si tomamos como espacio muestral

$$S = \{(c, c, c); (c, c, s); (c, s, c); (s, c, c); (c, s, s); (s, s, c); (s, c, s); (s, s, s)\}$$

entonces

$$X((c, c, c)) = 3$$

$$X((c, c, s)) = X((s, c, c)) = X((c, s, c)) = 2$$

$$X((c, s, s)) = X((s, c, s)) = X((s, s, c)) = 1$$

$$X((s, s, s)) = 0$$

La imagen de esta función es el conjunto $\{0, 1, 2, 3\}$

Dada una v.a. X a su imagen se la anota R_X y se la denomina **rango o recorrido de X**

En el ejemplo anterior $R_X = \{0, 1, 2, 3\}$

- 2- Se tira un dado tantas veces como sean necesarias hasta que sale el número 1 por primera vez.

Podemos simbolizar el espacio muestral de la siguiente manera $S = \{1, 01, 001, 0001, \dots\}$, por ejemplo 001 simboliza el resultado que en los dos primeros tiros no salió el número 1 y en el tercer tiro salió el 1.

Sea Y la v.a.:

Y : “número de tiros necesarios hasta que sale el 1 por primera vez”

Entonces $R_Y = \{1, 2, 3, 4, \dots\}$, es decir el rango de Y es el conjunto de los números naturales.

- 3- En el interior de un círculo de radio r y centro el origen de coordenadas, se elige un punto al azar.

Tomamos como espacio muestral a $S = \{(x, y), x^2 + y^2 \leq r^2\}$. Aquí S es infinito no numerable

Definimos la v.a. Z : “distancia del punto elegido al origen”

Entonces $R_Z = \{z; 0 \leq z \leq r\}$

Las variables aleatorias se clasifican según su rango.

Sea X es una v.a. con rango R_X . Si R_X es un conjunto *finito o infinito numerable* entonces se dice que X es una v.a. *discreta*. Si R_X es un conjunto *infinito no numerable* entonces X es una v.a. *continua*.

El rango R_X es considerado un nuevo *espacio muestral*, y sus subconjuntos son *eventos*.

Por ejemplo:

En el ejemplo 1, los eventos unitarios o elementales son $\{0\}, \{1\}, \{2\}, \{3\}$, *pero los anotamos* $\{X = 0\}, \{X = 1\}, \{X = 2\}, \{X = 3\}$

Otros eventos son, por ejemplo:

$\{X \leq 1\}$, es decir, salió a lo sumo una cara.

Notar que podemos escribir $\{X \leq 1\} = \{X = 0\} \cup \{X = 1\}$, o sea escribimos al evento como *unión de eventos elementales*

$\{X > 0\}$, es decir, salieron una o más caras. Tenemos que $\{X > 0\} = R_X - \{X = 0\}$

En el ejemplo 2, $\{Y \geq 4\}$ sería el evento “al menos 4 tiros son necesarios para que salga por primera vez el numero 1”

$\{4 \leq Y \leq 6\}$ sería el evento “se necesitan entre 4 y 6 tiros para que salga el 1 por primera vez”

Notar que $\{4 \leq Y \leq 6\} = \{Y = 4\} \cup \{Y = 5\} \cup \{Y = 6\}$

En el ejemplo 3, $\left\{\frac{1}{3}r < Z < \frac{2}{3}r\right\}$ sería el evento “el punto elegido se encuentra a una distancia del centro mayor que $\frac{1}{3}r$, pero menor que $\frac{2}{3}r$ ”

Volviendo al ejemplo 1, notar que $B = \{X = 0\}$ ocurre en R_X si y solo si el evento $A = \{(s, s, s)\}$ ocurre en S . Se dice que A y B son *eventos equivalentes*.

De la misma forma los eventos

$A = \{(c, c, c)\}$ y $B = \{X = 3\}$ son equivalentes

$A = \{(c, c, s), (c, s, c), (s, c, c)\}$ y $B = \{X = 2\}$ son equivalentes

En general

siendo $A \subset S$ y $B \subset R_X$, A y B son *equivalentes* si $A = \{s \in S; X(s) \in B\}$

Si X es una v.a. de S en R , y R_X es el rango de X , para calcular la probabilidad de un evento B de R_X se busca el evento A en S equivalente a B y entonces $P(B) = P(A)$

Por ejemplo,

En el ejemplo 1, $P(B) = P(X = 0) = P(A) = P(\{(s, s, s)\}) = \frac{1}{8}$ si la moneda es normal.

$P(B) = P(X \leq 1) = P(\{(s, s, s); (c, s, s); (s, c, s); (s, s, c)\}) = \frac{4}{8} = 0.5$ si la moneda es normal

También podríamos haber planteado

$$P(B) = P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{8} + \frac{3}{8} = 0.5$$

En el ejemplo 3, si $B = \left\{ \frac{1}{3}r \leq Z \leq \frac{2}{3}r \right\}$, entonces B es equivalente a

$A = \left\{ (x, y); \frac{1}{3}r \leq \sqrt{x^2 + y^2} \leq \frac{2}{3}r \right\}$, por lo tanto

$$P(B) = P(A) = \frac{\text{area de } A}{\text{area de } S} = \frac{\pi \left(\frac{2}{3}r \right)^2 - \pi \left(\frac{1}{3}r \right)^2}{\pi r^2} = \left(\frac{2}{3} \right)^2 - \left(\frac{1}{3} \right)^2 = \frac{3}{9} = \frac{1}{3}$$

Observación: en este ejemplo si $B = \left\{ Z = \frac{2}{3}r \right\}$, entonces $P(B) = P(A) = \frac{\text{area de } A}{\text{area de } S} = \frac{0}{\pi r^2} = 0$

3.2 - Variables aleatorias discretas

Sea X una v.a. discreta. Anotamos su rango como $R_X = \{x_1, x_2, \dots, x_n\}$ si el rango es un conjunto finito de n elementos, y anotamos $R_X = \{x_1, x_2, \dots\}$ si el rango es un conjunto infinito numerable.

A cada x_i se le asigna un número $p(x_i) = P(X = x_i)$. Estos números deben satisfacer las condiciones siguientes

a) $p(x_i) \geq 0$ para todo i

b) $\sum_i p(x_i) = 1$

La función $p(x)$ que antes se definió, se llama **función de probabilidad o de frecuencia de la v.a. X** . El conjunto de pares $(x_i, p(x_i))$ $i = 1, 2, \dots$ es la **distribución de probabilidad de X** .

Por ejemplo

1-Se tira una moneda normal tres veces, sea la v.a. X : “número de caras obtenidas”

Entonces $R_X = \{0, 1, 2, 3\}$

Para hallar la distribución de probabilidad de X supongamos que la probabilidad de salir cara es $\frac{1}{2}$ entonces

$$P(X = 0) = P(\{(s, s, s)\}) = \frac{1}{8}$$

$$P(X = 1) = P(\{(c, s, s); (s, c, s); (s, s, c)\}) = \frac{3}{8}$$

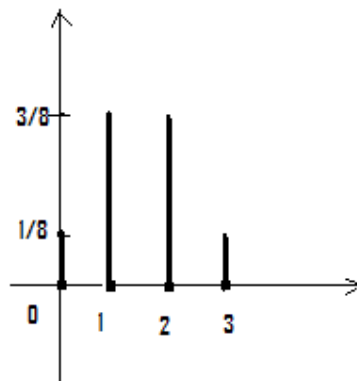
$$P(X = 2) = P(\{(c, c, s); (s, c, c); (c, s, c)\}) = \frac{3}{8}$$

$$P(X = 3) = P(\{(c, c, c)\}) = \frac{1}{8}$$

Se puede presentar la distribución de probabilidad de X en una tabla de la siguiente forma

x	0	1	2	3
$p(x)$	1/8	3/8	3/8	1/8

Un gráfico de la distribución de probabilidad de X sería



2-Se tira un dado normal. Sea X : “número que queda en la cara superior” Entonces $R_X = \{1, 2, 3, 4, 5, 6\}$
La función de distribución de X es

x	1	2	3	4	5	6
$p(x)$	1/6	1/6	1/6	1/6	1/6	1/6

Observación:

Sea X una v.a. discreta con rango finito $R_X = \{x_1, x_2, \dots, x_n\}$, donde cada x_i es un número entero y x_{i+1} es el consecutivo de x_i .

Si $P(X = x_i) = \frac{1}{n}$ para cada i entonces se dice que X tiene **distribución uniforme discreta**.

Por ejemplo podría ser $R_X = \{1, 2, 3, \dots, n-1, n\}$, en este caso X es uniforme discreta en el intervalo natural $[1, n]$.

La v.a. del ejemplo 2 anterior es uniforme discreta en el intervalo natural $[1, 6]$

Función de distribución acumulada

Sea X una v.a. con rango R_X . Se define la **función de distribución acumulada de X** (abreviamos F.d.a de X) como

$$F(x) = P(X \leq x) \quad -\infty < x < \infty \quad (12)$$

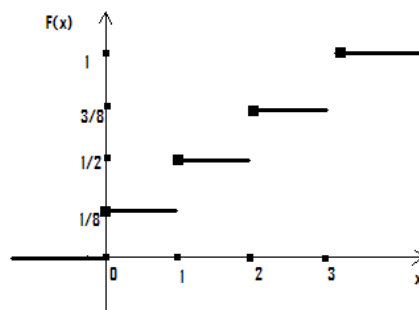
En el caso de ser X una v.a. discreta

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} p(x_i) \quad -\infty < x < \infty$$

Volviendo al ejemplo 1 anterior, la F.d.a. de X es

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{8} + \frac{3}{8} & \text{si } 1 \leq x < 2 \\ \frac{1}{8} + \frac{3}{8} + \frac{3}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x > 3 \end{cases} \Rightarrow F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{8} & \text{si } 0 \leq x < 1 \\ \frac{1}{2} & \text{si } 1 \leq x < 2 \\ \frac{7}{8} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x > 3 \end{cases}$$

La gráfica de la F.d.a. de X es



Observación: la F.d.a. de X es una función *escalonada*, los puntos de “salto” coinciden con los puntos del rango de X , y la magnitud del salto en x_i es igual a $P(X = x_i)$

En general si X es una v.a. discreta cualquiera, su F.d.a. será una función escalonada.

Además si x_1, x_2, \dots, x_n son los valores del rango de X ordenados de menor a mayor entonces

$$\begin{aligned} P(X = x_1) &= F(x_1) \\ P(X = x_i) &= F(x_i) - F(x_{i-1}) \quad i = 2, \dots, n \end{aligned}$$

Es decir, se puede obtener la función de distribución de X a partir de su F.d.a.

Para números cualesquiera a y b

1- Si $a \leq b$ entonces $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$

2- Si $a < b$ entonces $P(a < X \leq b) = P(X \leq b) - P(X \leq a)$

3- Si $a < b$ entonces $F(a) \leq F(b)$ (es decir $F(x)$ es una función creciente)

Además se cumple que

$$1 - \lim_{x \rightarrow \infty} F(x) = \lim_{x \rightarrow \infty} \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i} P(X = x_i) = 1$$

$$2 - \lim_{x \rightarrow -\infty} F(x) = \lim_{x \rightarrow -\infty} \sum_{x_i \leq x} P(X = x_i) = \sum_{x_i \leq -\infty} P(X = x_i) = 0$$

3.3 – Esperanza de una variable aleatoria discreta

Sea X una v.a. discreta con rango R_X . La **esperanza**, **valor medio** o **valor esperado de X** , lo anotamos $E(X)$, y se define como

$$E(X) = \sum_{x_i \in R_X} x_i P(X = x_i)$$

La sumatoria se hace sobre todos los posibles valores de X

Otra notación usual es μ_X o μ

Ejemplos:

1- Sea la v.a. X : “número que queda en la cara de arriba al tirar un dado normal”

$$R_X = \{1, 2, 3, 4, 5, 6\}$$

$$\begin{aligned} \text{Entonces } E(X) &= \sum_{x=1}^6 xP(X = x) = \\ &= 1P(X = 1) + 2P(X = 2) + 3P(X = 3) + 4P(X = 4) + 5P(X = 5) + 6P(X = 6) = \\ &= 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = \frac{7}{2} = 3.5 \end{aligned}$$

2- Se tira una moneda normal tres veces, sea la v.a. X : “número de caras obtenidas”

$$\text{Entonces } R_X = \{0, 1, 2, 3\}$$

Calculamos la esperanza de X

$$E(X) = \sum_{x=0}^3 xP(X = x) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2} = 1.5$$

Observaciones:

- 1- La esperanza de una v.a. no tiene que coincidir necesariamente con algún valor del rango de la variable
- 2- En el ejemplo 1 donde el rango es finito y equiprobable, la esperanza de X coincide con el **promedio de los valores del rango de X**
- 3- Se puede interpretar a la esperanza de una v.a. como un **promedio “pesado” o “ponderado” de los valores del rango de la variable**, donde el “peso” de cada x_i es la probabilidad $P(X = x_i)$
- 4- Otra interpretación que se puede hacer de la esperanza es la siguiente: consideremos el ejemplo 1, supongamos que tiramos el dado muchas veces, N veces, y entonces obtenemos una secuencia de N valores x_1, x_2, \dots, x_N donde cada x_i es un número natural del 1 al 6. Supongamos además que hacemos un promedio de esos N valores, y si llamamos n_i al número de veces que sale el número i tenemos que

$$\frac{x_1 + x_2 + \dots + x_N}{N} = \frac{n_1 1 + n_2 2 + \dots + n_6 6}{N} =$$

$$= \frac{n_1}{N} \times 1 + \frac{n_2}{N} \times 2 + \dots + \frac{n_6}{N} \times 6 \approx 1 \times P(X = 1) + 2 \times P(X = 2) + \dots + 6 \times P(X = 6) = E(X)$$

Es decir si promediamos los N valores medidos de X , ese promedio tiende a $E(X)$ cuando $N \rightarrow \infty$,
pues $\frac{n_i}{N} \approx P(X = i)$ cuando N es grande.

Esperanza de una función

A veces importa hallar la esperanza de una **función de X** y no de X misma. Veamos un ejemplo.
Un instructor de escritura técnica ha solicitado que cierto reporte sea entregado a la semana siguiente, agregando la restricción de que cualquier reporte que sobrepase las cuatro páginas será rechazado.
Sea X : “número de páginas del reporte de cierto estudiante seleccionado al azar”
Supongamos que X tenga la siguiente distribución de probabilidad

x	1	2	3	4
$p(x)$	0.01	0.19	0.35	0.45

Suponga que el instructor tarda \sqrt{X} minutos calificando un trabajo que consiste en X páginas. Claramente \sqrt{X} es **otra variable aleatoria**. ¿Cuál será su esperanza?, es decir ¿a qué es igual $E(\sqrt{X})$?
Para calcular la esperanza de una v.a. se necesita conocer su función de distribución de probabilidad, por lo tanto habría que hallar previamente la distribución de probabilidad de la v.a. $Y = \sqrt{X}$.
Está claro que si el rango de X es $R_X = \{1, 2, 3, 4\}$ entonces el rango de Y será $R_Y = \{\sqrt{1}, \sqrt{2}, \sqrt{3}, \sqrt{4}\}$.
Además

$$P(Y = \sqrt{1}) = P(X = 1) = 0.01$$

$$P(Y = \sqrt{2}) = P(X = 2) = 0.19$$

$$P(Y = \sqrt{3}) = P(X = 3) = 0.35$$

$$P(Y = \sqrt{4}) = P(X = 4) = 0.45$$

Por lo tanto

$$E(Y) = \sqrt{1} \times P(Y = \sqrt{1}) + \sqrt{2} \times P(Y = \sqrt{2}) + \sqrt{3} \times P(Y = \sqrt{3}) + \sqrt{4} \times P(Y = \sqrt{4}) =$$

$$= \sqrt{1} \times P(X = 1) + \sqrt{2} \times P(X = 2) + \sqrt{3} \times P(X = 3) + \sqrt{4} \times P(X = 4) = 1.78491$$

O sea

$$E(Y) = \sum_x \sqrt{x} P(X = x)$$

Lo visto en este ejemplo se puede generalizar en el siguiente

Teorema: Si X es una v.a. discreta con rango R_X y distribución de probabilidad $p(x)$, entonces la esperanza de cualquier función $h(X)$ es igual a

$$E(h(X)) = \sum_x h(x) p(x)$$

Ejemplo:

Un negocio de computadoras ha comprado tres computadoras de cierto tipo a \$500 cada una y las ven-

derá a \$1000 cada una. El fabricante ha aceptado volver a comprar en \$200 cualquier computadora que no se haya vendido en un tiempo especificado.

Sea X : “número de computadoras vendidas”, y supongamos que la distribución de probabilidad de X es

x	0	1	2	3
$p(x)$	0.1	0.2	0.3	0.4

Si consideramos la v.a. Y : “utilidad obtenida”, entonces Y es una función de X , es decir $Y = h(X)$

Específicamente

$$Y = 1000X + 200(3 - X) - 1500 = 800X - 900$$

La utilidad esperada, es decir la $E(Y)$ será

$$\begin{aligned} E(Y) &= \sum_{x=0}^3 (800x - 900)P(X = x) = \\ &= (800 \times 0 - 900)P(X = 0) + (800 \times 1 - 900)P(X = 1) + (800 \times 2 - 900)P(X = 2) + (800 \times 3 - 900)P(X = 3) = \\ &= (-900) \times 0.1 + (-100) \times 0.2 + 700 \times 0.3 + 1500 \times 0.4 = \$700 \end{aligned}$$

Notar que aplicando propiedades de la notación Σ se puede plantear

$$E(Y) = \sum_{x=0}^3 (800x - 900)P(X = x) = 800 \sum_{x=0}^3 xP(X = x) - 900 \sum_{x=0}^3 P(X = x) = 800E(X) - 900$$

y calculando la esperanza de X , se llega al mismo resultado

Propiedades de la esperanza

En el ejemplo anterior tenemos que Y es una **función lineal de X** , es decir $Y = aX + b$ con a y b números reales.

En este caso vale entonces la siguiente propiedad

$$E(aX + b) = aE(X) + b$$

La demostración sigue los mismos pasos que en el ejemplo anterior

$$\begin{aligned} E(aX + b) &= \sum_x (ax + b)P(X = x) = a \underbrace{\sum_x xP(X = x)}_{= E(X)} + b \underbrace{\sum_x P(X = x)}_{= 1} = aE(X) + b \end{aligned}$$

Ejemplo:

En el ejemplo anterior donde $Y = 800X - 900$

Directamente calculamos

$$E(Y) = 800E(X) - 900$$

Y

$$E(X) = 0 \times 0.1 + 1 \times 0.2 + 2 \times 0.3 + 3 \times 0.4 = 2$$

En consecuencia

$$E(Y) = 800E(X) - 900 = 800 \times 2 - 900 = 700$$

Observaciones:

- 1- Para cualquier constante a , $E(aX) = aE(X)$
- 2- Para cualquier constante b , $E(X + b) = E(X) + b$

Varianza de una variable aleatoria

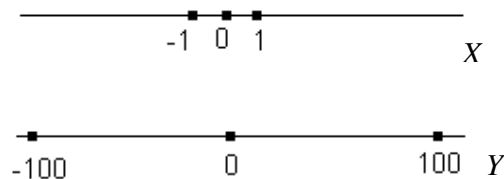
La esperanza de una v.a. mide dónde está centrada la distribución de probabilidad. Pero supongamos el siguiente ejemplo

Sean X e Y dos variables aleatorias con distribuciones dadas por

x	-1	1
$p(x)$	0.5	0.5

y	-100	100
$p(y)$	0.5	0.5

Es fácil verificar que $E(X) = E(Y) = 0$, pero los valores que toma la v.a. Y están más “alejados” de su esperanza que los valores de X .



Se busca una medida que refleje este hecho, se define entonces la varianza de una v.a.

Sea X una v.a. discreta con rango R_X , función de distribución de probabilidad $p(x)$ y esperanza $E(X) = \mu$,

Entonces la **varianza de X** , que anotamos $V(X)$, σ^2 o σ_X^2 es

$$V(X) = E[(X - \mu)^2] = \sum_{x \in R_X} (x - \mu)^2 p(x)$$

La **desviación estándar de X** es $\sigma_X = \sqrt{V(X)}$

Observaciones:

- 1- La varianza de una v.a. nunca es negativa
- 2- La cantidad $h(X) = (X - \mu)^2$ es el cuadrado de la desviación de X desde su media, y la varianza de X es la esperanza de la desviación al cuadrado. Si la mayor parte de la distribución de probabilidad está cerca de μ , entonces σ^2 será relativamente pequeña. Si hay valores de la variable alejados de μ que tengan alta probabilidad, entonces σ^2 será grande.
- 3- σ^2 está expresado en las unidades de medida de X al cuadrado, mientras que σ está expresada en las mismas unidades de medida que X .

Ejemplo:

En el caso de las variables aleatorias X e Y nombradas anteriormente,

$$V(X) = (-1 - 0)^2 \times 0.5 + (1 - 0)^2 \times 0.5 = 1 \quad \text{y} \quad \sigma_X = 1$$

$$V(Y) = (-100 - 0)^2 \times 0.5 + (100 - 0)^2 \times 0.5 = 100^2 \quad \text{y} \quad \sigma_Y = 100$$

Otra forma de escribir la varianza de una v.a., que facilita los cálculos es

$$\begin{aligned} V(X) &= \sum_{x \in R_X} (x - \mu)^2 p(x) = \sum_{x \in R_X} (x^2 - 2\mu x + \mu^2) p(x) = \sum_{x \in R_X} x^2 p(x) - 2\mu \sum_{x \in R_X} xp(x) + \mu^2 \sum_{x \in R_X} p(x) = \\ &= E(X^2) - 2\mu E(X) + \mu^2 = E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 \end{aligned}$$

Por lo tanto

$$V(X) = E(X^2) - \mu^2$$

Propiedades de la varianza

Las propiedades de la varianza de una v.a. son consecuencia de las propiedades de la esperanza de una v.a.

Si X es una v.a. discreta con rango R_X y distribución de probabilidad $p(x)$, entonces la varianza de cualquier función $h(X)$ es igual a

$$V(h(X)) = \sum_{x \in R_X} (h(x) - E(h(X)))^2 p(x)$$

Si $h(X)$ es una función lineal, entonces

$$V(aX + b) = a^2 V(X) \quad \text{y} \quad \sigma_{aX+b} = \sqrt{V(aX + b)} = |a| \sigma_X$$

Observaciones:

$$1- V(aX) = a^2 V(X)$$

$$2- V(X + b) = V(X)$$

Ejemplo:

En un ejemplo anterior donde X : “número de computadoras vendidas” y Y : “utilidad obtenida”, la $V(Y)$ sería $V(Y) = 800^2 V(X)$

Necesitamos calcular $V(X)$

$$V(X) = E(X^2) - \mu^2$$

Sabemos ya que $\mu = E(X) = 2$

$$\text{Calculamos } E(X^2) = 0^2 \times 0.1 + 1^2 \times 0.2 + 2^2 \times 0.3 + 3^2 \times 0.4 = 5$$

En consecuencia

$$V(Y) = 800^2 V(X) = 800^2 (E(X^2) - \mu^2) = 800^2 (5 - 2^2) = 800$$

3.4- Variables aleatorias discretas importantes

Distribución binomial

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos $P(A) = p$.

Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:

- 1- se realizan n repeticiones *independientes* de ε , donde n se fija de antemano.
 - 2- las repeticiones son idénticas, y en cada repetición de ε observamos *si ocurre A o no ocurre A* (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
 - 3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p
- Se dice entonces que ε_0 es un *experimento binomial*

Ejemplos:

- 1- Se tira una moneda 4 veces en forma sucesiva e independiente, y observamos en cada tiro si sale cara o no sale cara.
Entonces este es un experimento binomial pues:
 ε sería el experimento “tirar una moneda”
 A sería el evento “sale cara”
 ε se repite en forma sucesiva e independiente $n = 4$ veces
 $P(A) = p$ es la misma en cada tiro.
- 2- Se tiene una urna con 15 bolillas blancas y 5 verdes. Se extraen al azar *con reemplazo* tres bolillas y se observa si la bolilla extraída es blanca.
Entonces este es un experimento binomial pues:
 ε sería el experimento “extraer al azar una bolilla de la urna”
 A sería el evento “se extrae bolilla blanca”
 ε se repite en forma sucesiva e independiente $n = 3$ veces
 $P(A) = \frac{15}{20} = \frac{3}{4}$ es la misma en cada extracción.
- 3- Si en el ejemplo anterior se extraen las bolillas *sin reemplazo* entonces el experimento no es binomial, pues *falla la independencia*:

Si anotamos A_i : “se extrae bolilla blanca en la i – ésima extracción”, entonces

$$P(A_1) = \frac{15}{20} \quad ; \quad P(A_2) = P(A_2 / A_1)P(A_1) + P(A_2 / A_1^c)P(A_1^c) = \frac{14}{19} \times \frac{15}{20} + \frac{15}{19} \times \frac{5}{20} = \frac{15}{20}$$

Pero $P(A_2) = \frac{15}{20} \neq P(A_2 / A_1) = \frac{14}{19}$ por lo tanto las extracciones no son independientes

Observación: si en la urna hubiese 1500 bolillas blancas y 500 verdes y se extraen dos bolillas al azar sin reemplazo, entonces

$$P(A_2) = \frac{15}{20} = 0.75 \approx P(A_2 / A_1) = \frac{1499}{1999} = 0.74987$$

Por lo tanto en estas condiciones podemos asumir que el experimento es binomial

La variable aleatoria binomial y su distribución

En la mayoría de los experimentos binomiales, interesa el número total de éxitos, más que saber exactamente cuáles repeticiones produjeron los éxitos

Sea la v.a. X : “número de éxitos en las n repeticiones de ε ”

Entonces se dice que X es una v.a. binomial

Veamos cuál es la distribución de probabilidad de X , para esto primero tomamos un caso concreto: el ejemplo 1 anterior en el que se tira una moneda 4 veces. Supongamos que la probabilidad de cara es $\frac{3}{4}$

Aquí el rango de X sería $R_X = \{0,1,2,3,4\}$

Para facilitar la notación escribimos A_i : "sale cara en el i – ésimo tiro" $i = 1,2,3,4$

Por lo tanto

$$P(X = 0) = P(A_1^c \cap A_2^c \cap A_3^c \cap A_4^c) \underset{\substack{\swarrow \\ \text{por independencia}}}{=} P(A_1^c)P(A_2^c)P(A_3^c)P(A_4^c) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4} = \left(\frac{1}{4}\right)^4$$

Para calcular la $P(X = 1)$ pensamos que hay cuatro casos posibles en los que se puede obtener exactamente una cara, que la cara salga en el 1º tiro, o en el 2º o en el 3º o en el 4º tiro. Notar que tenemos cuatro casos y eso **es igual a la cantidad de formas en que podemos elegir entre los 4 tiros uno de ellos en el cual sale cara, es decir tenemos** $\binom{4}{1} = \frac{4!}{1!3!} = 4$ **casos diferentes.**

$$\begin{aligned} P(X = 1) &= \\ &= P(A_1 \cap A_2^c \cap A_3^c \cap A_4^c) + P(A_1^c \cap A_2 \cap A_3^c \cap A_4^c) + P(A_1^c \cap A_2^c \cap A_3 \cap A_4^c) + \\ &+ P(A_1^c \cap A_2^c \cap A_3^c \cap A_4) \end{aligned}$$

Cada término es igual a $p(1-p)^3$ por lo tanto $P(X = 1) = 4p(1-p)^3$

Análogamente, para calcular $P(X = 2)$ tenemos $\binom{4}{2} = \frac{4!}{2!2!} = 6$ casos en los que salen exactamente dos caras, por lo tanto

$$P(X = 2) = P(A_1 \cap A_2 \cap A_3^c \cap A_4^c) + P(A_1 \cap A_2^c \cap A_3 \cap A_4^c) + \dots = \binom{4}{2} p^2 (1-p)^2$$

Pensando de la misma forma los otros casos se llega a

$$P(X = 3) = \binom{4}{3} p^3 (1-p) \quad ; \quad P(X = 4) = p^4$$

En general con un argumento análogo tenemos que $R_X = \{0,1,2,\dots,n\}$ y

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0,1,2,\dots,n$$

Notación: indicamos que X es una v.a. binomial **con parámetros n y p** con el símbolo $X \sim B(n, p)$

Dado que los números $P(X = k)$ corresponden a la distribución de una v.a., automáticamente cumplen

$$\text{que } \sum_{k=0}^n P(X = k) = 1$$

De todas formas se podría hacer una verificación algebraica utilizando la fórmula del binomio de Newton

$$\sum_{k=0}^n P(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1$$

Ejemplos:

- 1- En el ejemplo anterior en el que se tira una moneda 4 veces, calcular la probabilidad de obtener:
- exactamente una cara
 - al menos una cara
 - a lo sumo una cara

Solución:

- a) tenemos que la v.a. X: “número de caras obtenido” es $B(4, 0.25)$

$$\text{se pide } P(X = 1) = \binom{4}{1} \left(\frac{1}{4}\right)^1 \left(1 - \frac{1}{4}\right)^3 = 4 \times \frac{1}{4} \times \left(\frac{3}{4}\right)^3 = 0.421875$$

- b) la probabilidad de obtener **al menos una cara** es

$$P(X \geq 1) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) = \sum_{k=1}^4 \binom{4}{k} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{4-k}$$

Pero más fácil es hacer

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{4}{0} \left(\frac{1}{4}\right)^0 \left(1 - \frac{1}{4}\right)^{4-0} = 1 - \left(\frac{3}{4}\right)^4 = 1 - 0.421875 = 0.578125$$

- c) la probabilidad de obtener **a lo sumo una cara** es

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{4}{0} \left(\frac{1}{4}\right)^0 \left(1 - \frac{1}{4}\right)^{4-0} + \binom{4}{1} \left(\frac{1}{4}\right)^1 \left(1 - \frac{1}{4}\right)^{4-1} = 0.84375$$

Observación: si $X \sim B(n, p)$ para calcular $P(X \leq k)$ en general se debe hacer

$$P(X \leq k) = \sum_{i=0}^k P(X = i) = P(X = 0) + P(X = 1) + \dots + P(X = k)$$

Notar que $P(X \leq k)$ es la F.d.a. de X evaluada en k, es decir $F(k) = P(X \leq k)$

Existen tablas de la función de distribución acumulada de la binomial para diferentes valores de n y p

Consultando estas tablas se puede obtener directamente el resultado del inciso c) buscando para $n = 4$ y $p = 0.25$

Además consultando las tablas podemos evaluar $P(X = k)$ haciendo

$$P(X = k) = F(k) - F(k - 1) \quad k = 1, 2, \dots, n$$

- 2- Supongamos que el 20% de todos los ejemplares de un texto en particular fallan en una prueba de resistencia a la encuadernación. Se seleccionan 15 ejemplares al azar.

Sea la v.a. X: “número de ejemplares que fallan en la prueba entre los 15 seleccionados”

- ¿cuál es la probabilidad de que a lo sumo 8 fallen en la prueba?
- ¿cuál es la probabilidad de que exactamente 8 fallen en la prueba?

c) ¿cuál es la probabilidad de que al menos 8 fallen en la prueba?

Solución:

a) Tenemos que $X \sim B(15, 0.2)$

$$P(X \leq 8) = \sum_{k=0}^8 P(X = k) = F(8) = 0.999$$

por tabla de la F.d.a.

$$b) P(X = 8) = F(8) - F(7) = 0.999 - 0.996 = 0.003$$

por tabla de la F.d.a.

$$c) P(X \geq 8) = 1 - P(X \leq 7) = 1 - F(7) = 1 - 0.996 = 0.004$$

por tabla de la F.d.a.

Observaciones:

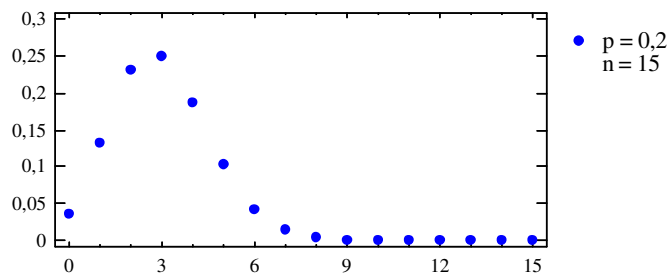
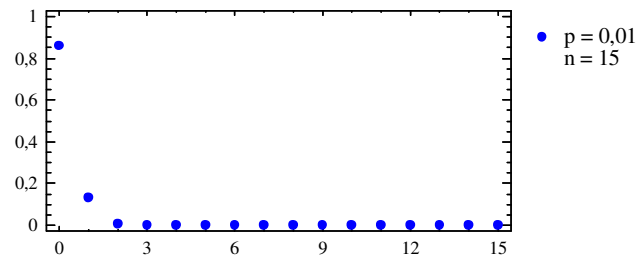
1- Si $X \sim B(1, p)$ entonces la v.a. X toma sólo dos valores 0 y 1 con probabilidades p y $1-p$ es decir podemos escribir

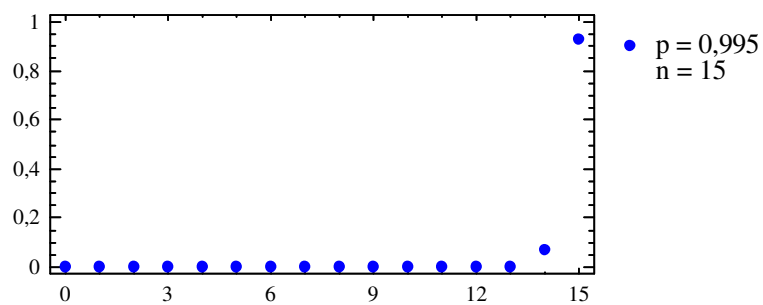
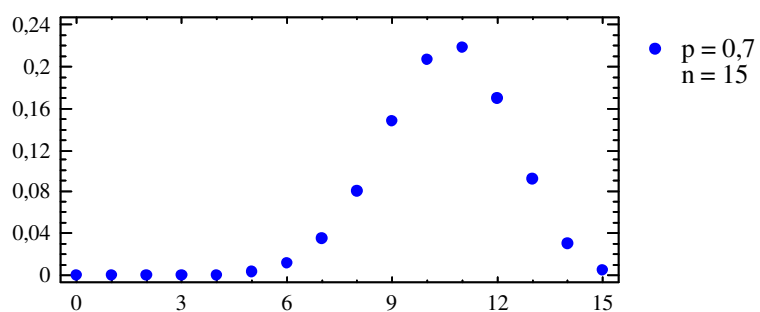
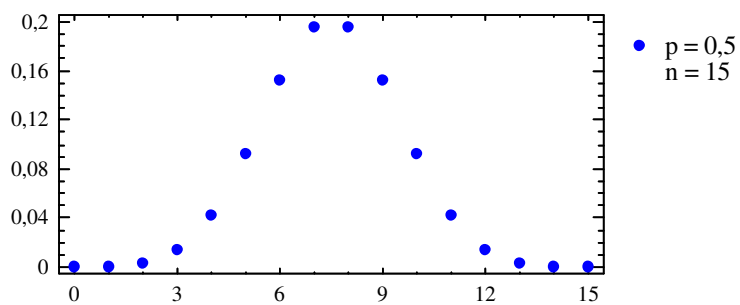
$$X = \begin{cases} 1 & \text{si al ejecutar } \varepsilon \text{ ocurre éxito} \\ 0 & \text{caso contrario} \end{cases} \quad \begin{aligned} P(X = 1) &= P(A) = p \\ P(X = 0) &= P(A^c) = 1 - p \end{aligned}$$

En este caso se dice que X tiene distribución de Bernoulli

En el caso de ser $X \sim B(n, p)$ se dice que se tienen “*n ensayos de Bernoulli*”

2- A continuación se muestra cómo varía la forma de la distribución a medida que p aumenta manteniendo n fijo en 15. Se grafica la distribución de frecuencia para $p = 0.01; 0.2, 0.5, 0.7$ y 0.995 . Observar que para $p = 0.5$ la distribución de frecuencia es simétrica.





Esperanza y varianza

Sea $X \sim B(n, p)$, entonces $E(X) = np$ y $V(X) = np(1 - p)$

Dem.)

Aplicamos la definición:

$$E(X) = \sum_{k=0}^n kp(k) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

El primer término es cero, por lo tanto podemos comenzar la suma en $k=1$:

$$E(X) = \sum_{k=1}^n k \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=1}^n \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k}$$

Ponemos todo en términos de $k-1$:

$$E(X) = \sum_{(k-1)=0}^n \frac{n(n-1)!}{(k-1)![(n-1)-(k-1)]!} p p^{k-1} (1-p)^{[(n-1)-(k-1)]}$$

Sacamos fuera de la suma n y p que no dependen del índice k y hacemos el cambio de índice: $(k-1) \rightarrow s$:

$$E(X) = np \sum_{s=0}^{n-1} \frac{(n-1)!}{s![(n-1)-s]!} p^s (1-p)^{[(n-1)-s]} = np \sum_{s=0}^{n-1} \binom{n-1}{s} p^s (1-p)^{[(n-1)-s]}$$

Recordando el desarrollo del binomio de Newton

$$(a+b)^r = \sum_{s=0}^{r} \binom{r}{s} a^s b^{r-s}, \text{ tenemos (con } r=n-1\text{):}$$

$$E(X) = np[p + (1-p)]^{n-1} = np[1]^{n-1} = np$$

Veamos el cálculo de la varianza

$$V(X) = E(X^2) - [E(X)]^2.$$

Luego:

$$V(X) = E(X^2) - [np]^2. \text{ Nos queda calcular } E(X^2).$$

$$E(X^2) = \sum_{k=0}^n k^2 \cdot p(k) = \sum_{k=0}^n k^2 \cdot \binom{n}{k} \cdot p^k (1-p)^{n-k} = \sum_{k=0}^n k^2 \cdot \frac{n!}{k!(n-k)!} \cdot p^k (1-p)^{n-k}.$$

Para calcular esta suma tratamos de llevarla a la forma del desarrollo de un binomio de Newton. Como el primer término es cero comenzamos a sumar desde $k=1$. Además simplificamos k en el numerador con el factor k en $k!$ del denominador:

$$E(X^2) = \sum_{k=1}^n k \cdot \frac{n!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k} = \sum_{k=1}^n k \cdot \frac{n(n-1)!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k}.$$

Separamos en dos sumas:

$$\begin{aligned} E(X^2) &= \sum_{k=1}^n [1 + (k-1)] \cdot \frac{n(n-1)!}{(k-1)!(n-k)!} \cdot p^k (1-p)^{n-k} \\ &= np \sum_{(k-1)=0}^{n-1} \frac{(n-1)!}{(k-1)![(n-1)-(k-1)]!} \cdot p^{(k-1)} (1-p)^{[(n-1)-(k-1)]} + \\ &\quad + n(n-1)p^2 \sum_{(k-2)=0}^{n-2} \frac{(n-2)!}{(k-1)![(n-2)-(k-2)]!} \cdot p^{(k-2)} (1-p)^{[(n-2)-(k-2)]} \end{aligned}$$

Esto es:

$$E(X^2) = np \sum_{s=0}^{n-1} k \cdot \binom{n-1}{s} \cdot p^s (1-p)^{[(n-1)-s]} + n(n-1)p^2 \sum_{r=0}^{n-2} k \cdot \binom{n-2}{r} \cdot p^r (1-p)^{[(n-2)-r]}$$

Las sumas corresponden al desarrollo de un binomio de Newton:

$$E(X^2) = np[p + (1-p)]^{n-1} + n(n-1)p^2[p + (1-p)]^{n-2} = np[1]^{n-1} + n(n-1)p^2[1]^{n-2}, \text{ es decir}$$

$$E(X^2) = np + n(n-1)p^2. \text{ Entonces:}$$

$$V(X) = E(X^2) - [np]^2 = np + n(n-1)p^2 - [np]^2 = np - np^2 \text{ o sea:}$$

$$V(X) = np(1-p)$$

Distribución geométrica

Sea \mathcal{E} un experimento aleatorio. Sea A un evento asociado a \mathcal{E} y anotamos $P(A) = p$.

Supongamos un experimento aleatorio \mathcal{E}_0 que cumple los siguientes requisitos:

- 1- se realizan repeticiones independientes de \mathcal{E} , hasta que ocurre A por primera vez inclusive.
- 2- las repeticiones son idénticas, y en cada repetición de \mathcal{E} observamos **si ocurre A o no ocurre A** (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
- 3- la probabilidad de éxito es constante de una repetición a otra de \mathcal{E} , y es igual a p

Sea la v.a. X : “número de repeticiones de \mathcal{E} hasta que ocurre A por primera vez inclusive”

Veamos cuál es la distribución de probabilidad de X

El rango es el conjunto $R_X = \{1, 2, 3, \dots\} = \mathbb{N}$

Además si anotamos A_i : “ocurre A en la i -ésima repetición de \mathcal{E} ”, entonces

$$P(X = k) = P(A_1^C \cap A_2^C \cap \dots \cap A_{k-1}^C \cap A_k) = P(A_1^C)P(A_2^C) \dots P(A_{k-1}^C)P(A_k) = (1-p)^{k-1} p$$

↓
por independencia

En consecuencia

$$P(X = k) = (1-p)^{k-1} p \quad k = 1, 2, \dots$$

Notación: $X \sim G(p)$

Para verificar que $\sum_{k=1}^{\infty} P(X = k) = 1$ recordar que $\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$ si $|a| < 1$

$$\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} (1-p)^{k-1} = p \frac{1}{1-(1-p)} = 1$$

Observaciones:

$$1- P(X = k) = (1-p)^{k-1} p \quad \text{y} \quad P(X = k+1) = (1-p)^k p$$

$$\text{Es decir } P(X = k+1) = (1-p)^{k-1} p(1-p) = P(X = k)(1-p) \quad \forall k = 1, 2, \dots$$

Podemos interpretar que para cada k , los $a_k = P(X = k)$ son los términos de una sucesión geométrica con razón $1-p$ y primer término $a_1 = p$

2- La F.d.a. sería

$$F(x) = P(X \leq x) = \sum_{k=1}^{\lfloor x \rfloor} P(X = k)$$

donde $\lfloor x \rfloor$ indica parte entera de x

Si x es un entero positivo entonces recordando que la suma de los n primeros términos de una sucesión geométrica con razón r y término general $a_k = pr^{k-1}$ es $\sum_{k=1}^n a_k = \frac{a_1(1-r^n)}{1-r}$ tenemos que

$$F(x) = P(X \leq x) = \sum_{k=1}^{\lfloor x \rfloor} P(X = k) = \frac{p(1-r^{\lfloor x \rfloor})}{1-r} = \frac{p(1-(1-p)^{\lfloor x \rfloor})}{1-(1-p)} = 1-(1-p)^{\lfloor x \rfloor}$$

$r = 1-p$

Por lo tanto

$$F(x) = \begin{cases} 1-(1-p)^{\lfloor x \rfloor} & \text{si } x \geq 1 \\ 0 & \text{caso contrario} \end{cases}$$

3- Una variante en la definición de distribución geométrica es definir la v.a. Y : “**número de fracasos** hasta el primer éxito”, en este caso $R_Y = \{0,1,2,\dots\}$, es decir se incluye al cero.

La distribución de probabilidad o frecuencia sería en este caso

$$P(Y = k) = (1-p)^k p \quad k = 0,1,2,\dots$$

Notar que la relación entre X e Y sería: $X = Y + 1$.

Es decir si adoptamos esta última definición **no incluimos** la repetición del experimento en el cual ocurre el primer éxito.

Ejemplos:

1- La probabilidad de que una computadora que corre cierto sistema operativo se descomponga en determinado día es de 0.1. Determinar la probabilidad de que la máquina se descomponga por primera vez en el duodécimo día, después de la instalación del sistema operativo

Solución:

Definimos la v.a. X : “número de días hasta que la computadora se descompone por primera vez”

Entonces $X \sim G(0.1)$

Se pide calcular la $P(X = 12)$

$$P(X = 12) = (1-p)^{12-1} p = 0.9^{11} \times 0.1 = 0.031381$$

2- Una prueba de resistencia a la soldadura consiste en poner carga en uniones soldadas hasta que se dé una ruptura. Para cierto tipo de soldadura, 80% de las rupturas ocurre en la propia soldadura, mientras que otro 20% se da en las vigas. Se prueba cierto número de soldaduras.

Sea la v.a. X : “número de pruebas hasta que se produce la ruptura de la viga”

¿Qué distribución tiene X ? ¿Cuál es la probabilidad que en la tercera prueba se produzca la primera ruptura de la viga?

Solución:

Cada prueba es un “*ensayo de Bernoulli*”, con un éxito definido como la ruptura de una viga. Por lo tanto, la probabilidad de éxito es $p = 0.2$.

La v.a. X tiene una distribución geométrica con parámetro $p = 0.2$ es decir $X \sim G(0.2)$

Para calcular la probabilidad pedida hacemos $P(X = 3) = (1 - p)^{3-1} p = 0.8^2 \times 0.2 = 0.128$

Esperanza y varianza

Sea $X \sim G(p)$ entonces $E(X) = \frac{1}{p}$ y $V(X) = \frac{1-p}{p^2}$

Dem.) Llamamos $1 - p = q$

Planteamos

$$\sum_{k=1}^{\infty} kP(X = k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1} =$$

Notar que $\frac{d}{dq}(q^k) = kq^{k-1}$, por lo tanto como $\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$ si $|a| < 1$

$$= p \sum_{k=1}^{\infty} \frac{d}{dq}(q^k) = p \frac{d}{dq} \left(\sum_{k=1}^{\infty} q^k \right) = p \frac{d}{dq} \left(\frac{q}{1-q} \right) = p \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \quad \therefore E(X) = \frac{1}{p}$$

Calculamos ahora la varianza

$$V(X) = E(X^2) - \mu^2 \text{ donde}$$

$$\mu = E(X) = \frac{1}{p}$$

$$E(X^2) = \sum_{k=1}^{\infty} k^2 P(X = k) = \sum_{k=1}^{\infty} k^2 p(1-p)^{k-1} = \sum_{k=1}^{\infty} k(k-1+1)p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(k-1)q^{k-1} + \sum_{k=1}^{\infty} kpq^{k-1}$$

Pero

$$\begin{aligned} p \sum_{k=1}^{\infty} k(k-1)q^{k-1} &= pq \sum_{k=1}^{\infty} k(k-1)q^{k-2} = pq \sum_{k=1}^{\infty} \frac{d^2}{dq^2}(q^k) = pq \frac{d^2}{dq^2} \left(\sum_{k=1}^{\infty} q^k \right) = pq \frac{d^2}{dq^2} \left(\frac{q}{1-q} \right) = \\ &= pq \frac{2}{(1-q)^3} = q \frac{2}{p^2} \end{aligned}$$

$$\text{Y } \sum_{k=1}^{\infty} kpq^{k-1} = \frac{1}{p}$$

Por lo tanto

$$V(X) = E(X^2) - \frac{1}{p^2} = q \frac{2}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q + p - 1}{p^2} = \frac{2(1-p) + p - 1}{p^2} = \frac{1-p}{p^2}$$

Distribución binomial negativa

La distribución binomial negativa constituye una extensión de la distribución geométrica.

Sea r un entero positivo.

Sea ε un experimento aleatorio. Sea A un evento asociado a ε y anotamos $P(A) = p$.

Supongamos un experimento aleatorio ε_0 que cumple los siguientes requisitos:

- 1- se realizan repeticiones independientes de ε , hasta que ocurre A por r -ésima vez inclusive.
- 2- las repeticiones son idénticas, y en cada repetición de ε observamos **si ocurre A o no ocurre A** (cuando A ocurre se dice que se obtuvo un “éxito”, caso contrario se obtuvo un “fracaso”)
- 3- la probabilidad de éxito es constante de una repetición a otra de ε , y es igual a p

Sea la v.a. X : “número de repeticiones de ε hasta que ocurre A por r -ésima vez, incluyendo la r -ésima vez que ocurre A ”

Veamos cuál es la distribución de probabilidad de X

El rango es el conjunto $R_X = \{r, r+1, r+2, r+3, \dots\}$

Para obtener una expresión genérica de la $P(X = k)$, notar que si en el k -ésimo ensayo ocurre éxito por r -ésima vez, entonces en los $k-1$ primeros ensayos ocurrieron $r-1$ éxitos. Si anotamos B : “en los primeros $k-1$ ensayos ocurran $r-1$ éxitos”, entonces

$$P(B) = \binom{k-1}{r-1} p^{r-1} (1-p)^{(k-1)-(r-1)} = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r}$$

Además si anotamos A : “ocurre A en la r -ésima repetición de ε ”, entonces $P(A) = p$ y A y B son independientes, por lo tanto

$$P(X = k) = P(B \cap A) = P(B)P(A) = \binom{k-1}{r-1} p^{r-1} (1-p)^{k-r} \times p = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

En resumen

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \quad k = r, r+1, r+2, \dots$$

Notación: $X \sim BN(r, p)$

Para verificar que $\sum_{k=0}^n P(X = k) = 1$ se deriva r veces la igualdad $\sum_{k=0}^{\infty} (1-p)^{k-1} = \frac{1}{p}$ la que se deduce

$$\text{de } p \sum_{k=0}^{\infty} (1-p)^{k-1} = 1$$

Ejemplo:

En una prueba de fuerza de soldadura, 80% de las pruebas da como resultado ruptura de soldadura, mientras que otro 20% da ruptura de la viga. Sea la v.a. X : “número de pruebas hasta la tercera ruptura de la viga inclusive”. ¿Cuál es la distribución de X ? Determinar la $P(X = 8)$

Solución:

Tenemos que $X \sim BN(3, 0.2)$

$$\text{Por lo tanto } P(X = 8) = \binom{8-1}{3-1} p^3 (1-p)^{8-3} = \binom{7}{2} 0.2^3 \times 0.8^5 = 0.05505$$

Observación: la distribución geométrica puede verse como un caso particular de la distribución binomial negativa con $r = 1$

Esperanza y varianza

$$\text{Si } X \sim BN(r, p) \text{ entonces } E(X) = \frac{r}{p} \text{ y } V(X) = \frac{r(1-p)}{p^2}$$

Dem.) Se hará mas adelante

Distribución hipergeométrica

Supongamos que tenemos una población o conjunto de N objetos o individuos (es decir tenemos una población finita).

Clasificamos a los objetos de la población en dos categorías. Hay M objetos de una categoría y $N-M$ de la otra categoría. Se suele decir que tenemos M “éxitos” y $N-M$ “fracasos”.

Se extraen al azar y sin reemplazo n objetos de dicha población. Es decir se extrae una muestra de n objetos de la población, de manera tal que es igualmente probable que se seleccione cada subconjunto de tamaño n .

Consideramos la v.a. X : “número de éxitos en la muestra extraída”

Se dice que X tiene una distribución hipergeométrica con parámetros n , M y N

Notación: $X \sim H(n, M, N)$

Veamos cuál es la distribución de X

Primero notar que una expresión para la $P(X = k)$, usando combinatoria es

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad \text{donde para que los números combinatorios estén bien definidos debe}$$

cumplirse $0 \leq k \leq M$ y $0 \leq n-k \leq N-M$. Pero estas condiciones son equivalentes a

$$\max(0, n - N + M) \leq k \leq \min(n, M)$$

Por lo tanto la distribución de probabilidad de X es

$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} \quad \max(0, n - N + M) \leq k \leq \min(n, M)$$

Se verifica que $\sum_{k=0}^n P(X = k) = 1$ pues los números $P(X = k)$ corresponden a la distribución de una v.a.

Ejemplo:

- 1- De 50 edificios en un parque industrial, 12 no cumplen el código eléctrico. Si se seleccionan aleatoriamente 10 edificios para inspeccionarlos, ¿cuál es la probabilidad de que exactamente tres de los diez no cumplan el código?

Solución:

Sea la v.a. X : “número de edificios seleccionados que violan el código”, entonces $X \sim H(10, 12, 50)$. se pide calcular la $P(X = 3)$

$$P(X = 3) = \frac{\binom{12}{3} \binom{50-12}{10-3}}{\binom{50}{10}} = 0.2703$$

- 2- Un cargamento contiene 40 elementos. Se seleccionará de forma aleatoria y se probará 5 elementos. Si dos o más están defectuosos, se regresará el cargamento.
- a) si de hecho el cargamento contiene cinco elementos defectuosos, ¿cuál es la probabilidad de que sean aceptados?
- b) si de hecho el cargamento contiene diez elementos defectuosos, ¿cuál es la probabilidad de que no sean aceptados?

Solución:

- a) Sea la v.a. X : “número de elementos defectuosos en la muestra”

En este caso $X \sim H(5, 5, 40)$. Hay que calcular la $P(X \leq 1)$

$$P(X \leq 1) = P(X = 0) + P(X = 1)$$

$$P(X = 0) = \frac{\binom{5}{0} \binom{40-5}{5-0}}{\binom{40}{5}} = 0.4933557 \quad P(X = 1) = \frac{\binom{5}{1} \binom{40-5}{5-1}}{\binom{40}{5}} = 0.3978675$$

$$\therefore P(X \leq 1) = 0.4933557 + 0.3978675 = 0.8912232$$

- b) Sea la v.a. X : “número de elementos defectuosos en la muestra”

En este caso $X \sim H(5, 10, 40)$. Hay que calcular la $P(X \geq 2)$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1)$$

$$P(X = 0) = \frac{\binom{10}{0} \binom{40-10}{5-0}}{\binom{40}{5}} = 0.2165718 \quad P(X = 1) = \frac{\binom{10}{1} \binom{40-10}{5-1}}{\binom{40}{5}} = 0.416484$$

$$\therefore P(X \geq 2) = 0.3669442$$

Observación:

En el ejemplo anterior si el cargamento hubiese tenido 400 elementos se podría haber considerado en la parte a) a X con distribución binomial con parámetros $n = 5$ y $p = \frac{5}{400}$

En general, si el tamaño de la población N y el número de éxitos M crecen pero de manera tal que $\frac{M}{N} \rightarrow p$ y n es chico comparado con N , se puede verificar que

$$\lim_{N \rightarrow \infty} \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}} = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{donde } \frac{M}{N} = p$$

Por lo tanto, para una fracción fija de defectuosos $\frac{M}{N} = p$ la función de probabilidad hipergeométrica converge a la función de probabilidad binomial cuando N se hace grande.

Esperanza y varianza

$$\text{Si } X \sim H(n, M, N) \text{ entonces } E(X) = \frac{nM}{N} \text{ y } V(X) = n \frac{M}{N} \left(\frac{N-M}{N} \right) \left(\frac{N-n}{N-1} \right)$$

Dem.) La demostración se hará mas adelante.

Distribución de Poisson

Una v.a. X con rango $R_X = \{0, 1, 2, \dots\}$ se dice tener *distribución de Poisson con parámetro λ* , si para algún $\lambda > 0$

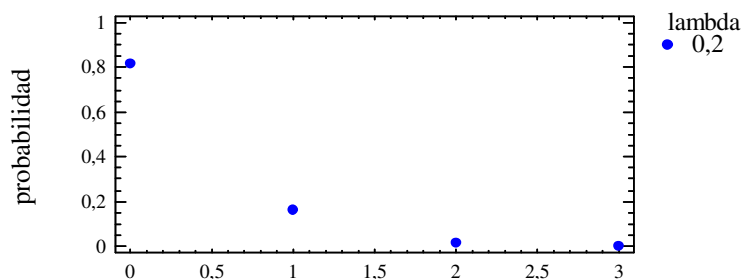
$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \quad k = 0, 1, 2, \dots$$

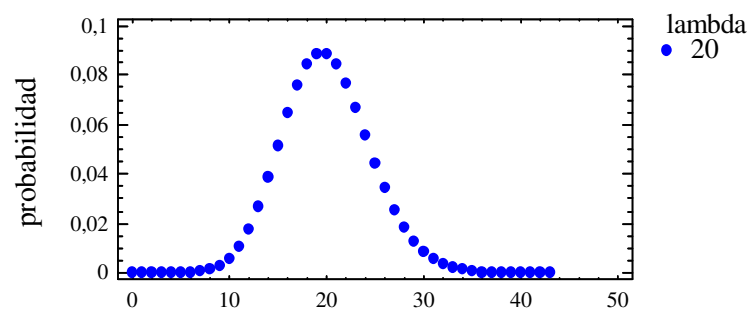
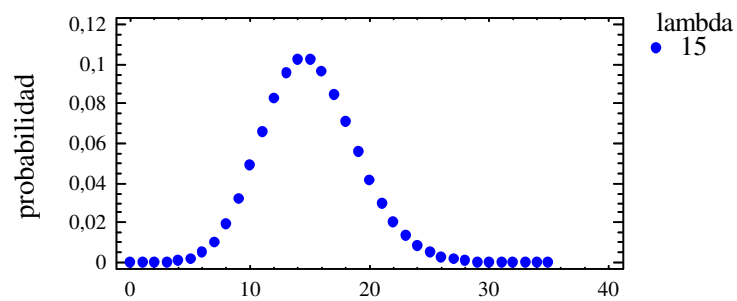
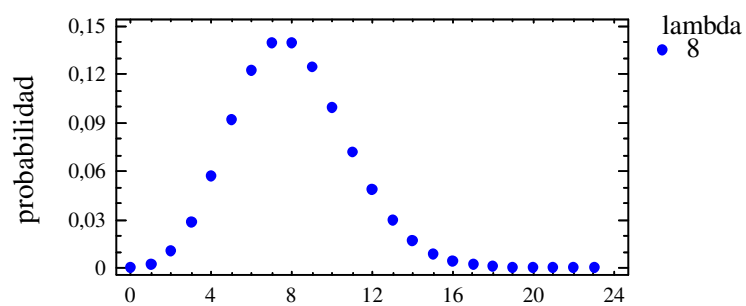
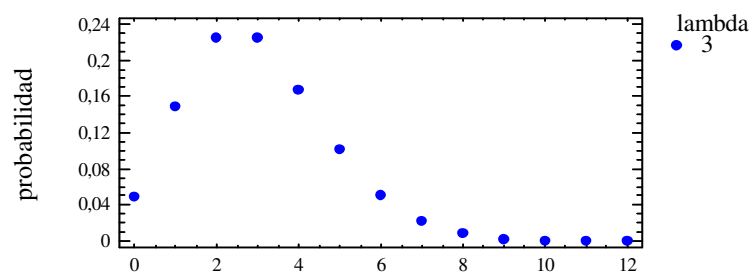
Es fácil verificar que $\sum_{k=0}^{\infty} P(X = k) = 1$ usando el hecho que $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$

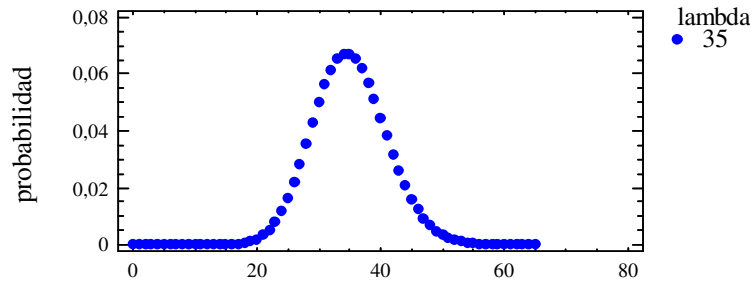
$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

En los siguientes gráficos se ve como varía la forma de la distribución con los valores de λ

Notar que para valores de λ “pequeños” la distribución es asimétrica, a medida que λ aumenta, la distribución tiende a ser cada vez más simétrica







Ejemplo:

Considere escribir en un disco de computadora y luego enviar el escrito por un certificador que cuenta el número de pulsos faltantes. Suponga que este número X tiene una distribución de Poisson con parámetro igual a 0.2

- ¿Cuál es la probabilidad de que un disco tenga exactamente un pulso faltante?
- ¿Cuál es la probabilidad de que un disco tenga al menos dos pulsos faltantes?
- Si dos discos se seleccionan independientemente, ¿cuál es la probabilidad de que ninguno contenga algún pulso faltante?

Solución:

a) Sea la v.a. X : “número de pulsos faltantes en un disco”

Entonces $X \sim P(0.2)$

$$\text{Se pide } P(X = 1) = e^{-0.2} \frac{0.2^1}{1!} = 0.163746$$

b) Siendo X como en a) se pide calcular $P(X \geq 2)$

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) = 1 - e^{-0.2} \frac{0.2^0}{0!} - e^{-0.2} \frac{0.2^1}{1!} = 0.01752$$

c) Sea la v.a. Y : “número de discos sin pulsos faltantes”

Entonces $Y \sim B(2, p)$ donde $p = P(X = 0) = e^{-0.2}$

Por lo tanto se pide calcular $P(Y = 2)$

$$P(Y = 2) = \binom{2}{2} p^2 (1 - p)^0 = p^2 = (e^{-0.2})^2 = 0.67032$$

Esperanza y varianza

Si $X \sim P(\lambda)$ entonces $E(X) = \lambda$ y $V(X) = \lambda$

Dem.)

$$E(X) = \sum_{k=0}^{\infty} k P(X = k) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \lambda \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^{k-1}}{(k-1)!} = \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda$$

$$V(X) = E(X^2) - \mu^2 \quad \text{donde } \mu = E(X) = \lambda$$

$$E(X^2) = \sum_{k=0}^{\infty} k^2 P(X = k) = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{\infty} k \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=1}^{\infty} (k-1+1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} =$$

$$\begin{aligned}
 &= \sum_{k=1}^{\infty} (k-1) \frac{e^{-\lambda} \lambda^k}{(k-1)!} + \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=2}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-2)!} + \sum_{k=1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{k+2}}{k!} + \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^{k+1}}{k!} = \\
 &= \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} + \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = \lambda^2 + \lambda
 \end{aligned}$$

Por lo tanto

$$V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Aplicaciones de la distribución de Poisson

La v.a. Poisson tiene un gran rango de aplicaciones, una de ellas es la aproximación para una v.a. binomial con parámetros n y p cuando n es grande y p es pequeño de manera tal que $np \rightarrow \lambda$, específicamente, sea $X \sim B(n, p)$ y sea $\lambda = np$, entonces

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n(n-1)\dots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1 - \frac{\lambda}{n}\right)^n$$

Para n grande y p chico

$$\left(1 - \frac{\lambda}{n}\right)^n \approx e^{-\lambda} ; \quad \frac{n(n-1)\dots(n-k+1)}{n^k} \approx 1 ; \quad \left(1 - \frac{\lambda}{n}\right)^k \approx 1$$

Entonces, para n grande y p chico

$$P(X = k) \approx e^{-\lambda} \frac{\lambda^k}{k!}$$

Es decir cuando n es grande, p chico y np es “moderado” entonces la v.a. binomial con parámetros n y p tiene una distribución que se **aproxima** a la de una Poisson con parámetro $\lambda = np$

Ejemplo:

Supongamos que la probabilidad de que un artículo producido por cierta máquina sea defectuoso es 0.1. Hallar la probabilidad que una muestra de 10 artículos contenga a lo sumo un defectuoso.

Sea X : “número de artículos defectuosos en la muestra”

Podemos asumir que $X \sim B(10, 0.1)$

La probabilidad **exacta** pedida es

$$P(X \leq 1) = P(X = 0) + P(X = 1) = \binom{10}{0} (0.1)^0 (0.9)^{10-0} + \binom{10}{1} (0.1)^1 (0.9)^{10-1} = 0.7361$$

La **aproximación** de Poisson da

$$P(X \leq 1) = P(X = 0) + P(X = 1) \approx e^{-1} \frac{1^0}{0!} + e^{-1} \frac{1^1}{1!} = e^{-1} + e^{-1} \approx 0.7358$$

$$\lambda = np = 10 \times 0.1 = 1$$

Algunos autores sostienen que la aproximación de Poisson funciona bien cuando n es grande, p es chico y $np < 7$ (Mendenhall, Estadística matemática con aplicaciones), otros recomiendan usar la aproximación de Poisson a la binomial cuando $n \geq 100$ y $p \leq 0.01$ y $np \leq 20$ (Devore, Probabilidad para ingeniería y ciencias)

En la siguiente tabla se da un ejemplo de una aproximación de Poisson a la función de distribución de la binomial. Se tabula la $P(X = k)$ para algunos valores de k para las distribuciones binomial y Poisson con los parámetros que se indican

k	$X = B(10^5, 2 \times 10^{-5})$	$X = B(5 \times 10^3, 4 \times 10^{-5})$	P(2)
0	0.135281	0.135281	0.135335
1	0.270671	0.270671	0.270671
2	0.270725	0.270725	0.270671
3	0.180483	0.180483	0.180447
4	0.0902235	0.0902235	0.0902235
5	0.036075	0.036075	0.0360894
6	0.0120178	0.0120178	0.0120298
7	0.0034309	0.0034309	0.00343709
8	0.000856867	0.000856867	0.000859272
9	0.000190186	0.000190186	0.000190949
10	0.000037984	0.000037984	0.0000381899
11	6.89513×10^{-6}	6.89513×10^{-6}	6.94361×10^{-6}
12	1.14712×10^{-6}	1.14712×10^{-6}	1.15727×10^{-6}
13	1.76127×10^{-7}	1.76127×10^{-7}	1.78041×10^{-7}
14	2.51056×10^{-8}	2.51056×10^{-8}	2.54345×10^{-8}
15	3.33937×10^{-9}	3.33937×10^{-9}	3.39126×10^{-9}

Ejemplo:

En una prueba de tarjetas de circuitos, la probabilidad de que un diodo en particular falle es 0.01. Suponga que una tarjeta contiene 200 diodos.

- ¿Cuál es la probabilidad aproximada de que por lo menos 4 diodos fallen en una tarjeta seleccionada al azar?
- Si se embarcan cinco tarjetas a un cliente en particular, ¿cuál es la probabilidad de que por lo menos cuatro de ellas funcionen bien? (Una tarjeta funciona bien solo si todos sus diodos funcionan bien)

Solución:

- Sea la v.a. X : “número de diodos en una tarjeta que fallan”

Entonces $X \sim B(200, 0.01)$. Como n es grande y p chico aplicamos la aproximación Poisson con $np = 200 \times 0.01 = 2$

Se pide calcular la $P(X \geq 4)$

$$P(X \geq 4) = 1 - P(X \leq 3) \approx 1 - 0.857 = 0.143$$

por tabla de la acumulada de la Poisson

- Sea la v.a. Y : “número de tarjetas entre 5 que funcionan bien”

Tenemos que $Y \sim B(5, p)$ donde $p = P(X = 0) \approx e^{-2} \frac{2^0}{0!} = e^{-2}$

Se pide calcular $P(Y \geq 4)$

$$P(Y \geq 4) = P(Y = 4) + P(Y = 5) = \binom{5}{4} (e^{-2})^4 (1 - e^{-2})^{5-4} + \binom{5}{5} (e^{-2})^5 (1 - e^{-2})^{5-5} = 5e^{-8} (1 - e^{-2}) + 5e^{-10}$$

Proceso de Poisson

Una aplicación importante de la distribución de Poisson se presenta en relación con el acontecimiento de eventos de un tipo particular en el tiempo. Por ejemplo, un evento podría ser un individuo entrando en un establecimiento en particular, o pulsos radiactivos registrados por un contador Geiger, o automóviles pasando por un cruce determinado.

Supongamos que tenemos eventos que ocurren en ciertos puntos aleatorios de tiempo, y asumimos que para alguna constante positiva λ las siguientes suposiciones se sostienen:

- 1- La probabilidad que exactamente 1 evento ocurra en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $\lambda t + o(t)$ (donde $o(t)$ simboliza una función $f(t)$ tal que $\lim_{t \rightarrow 0} \frac{f(t)}{t} = 0$, por ejemplo $f(t) = t^2$ es $o(t)$, pero $f(t) = t$ no lo es)
- 2- La probabilidad que 2 o mas eventos ocurran en un intervalo de longitud t es la misma para todos los intervalos de longitud t y es igual a $o(t)$.
- 3- Para cualesquiera enteros n, k_1, k_2, \dots, k_n y cualquier conjunto de n intervalos I_1, I_2, \dots, I_n que no se superpongan, si definimos los eventos E_i : "en el intervalo I_i ocurren exactamente k_i eventos" $i = 1, 2, \dots, n$, entonces los eventos E_1, E_2, \dots, E_n son independientes.

Las suposiciones 1 y 2 establecen que para pequeños valores de t , la probabilidad de que exactamente un evento ocurra en un intervalo de longitud t es igual a λt mas algo que es chico comparado con t , mientras que la probabilidad de que 2 o mas eventos ocurran es pequeño comparado con t . la suposición 3 establece que lo que ocurra en un intervalo no tiene efecto (en la probabilidad) sobre lo que ocurrirá en otro intervalo que no se superponga.

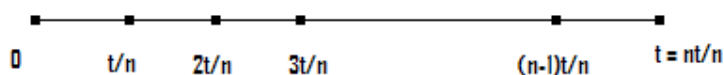
Bajo las suposiciones 1, 2 y 3 se puede probar que la v.a.

X : "número de eventos que ocurren en cualquier intervalo de longitud t " , tiene distribución Poisson con parámetro λt

Específicamente

$$P(X = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad k = 0, 1, 2, \dots$$

La idea de la demostración es la siguiente, partimos al intervalo $[0, t]$ en n subintervalos que no se superpongan cada uno de longitud $\frac{t}{n}$



Elegimos n suficientemente grande para que en cada subintervalo se tenga una ocurrencia exactamente o ninguna ocurrencia.

Sea la v.a. Y : “número de subintervalos en los que hay exactamente una ocurrencia”, entonces podemos asumir que $Y \sim B(n, p)$ donde p es la probabilidad que en un subintervalo hay exactamente una ocurrencia y si n es grande entonces la longitud del subintervalo $\frac{t}{n}$ es chica con lo cual por suposición

1 tenemos que $p \approx \lambda \frac{t}{n}$.

Entonces, utilizando la aproximación de Poisson a la binomial con parámetro $np = n\lambda \frac{t}{n} = \lambda t$ tenemos

$$P(X = k) = P(Y = k) \approx e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad k = 0, 1, 2, \dots$$

Observaciones:

- 1- Un **proceso temporal de Poisson** consiste en eventos que ocurren en el tiempo en forma aleatoria que cumplen con las suposiciones 1, 2 y 3.
- 2- El parámetro λ es la **tasa o rapidez del proceso**.
- 3- Si en lugar de observar eventos en el tiempo, consideramos observar eventos de algún tipo que ocurren en una región de dos o tres dimensiones, por ejemplo, podríamos seleccionar de un mapa una región R de un bosque, ir a esa región y contar el número de árboles. Cada árbol representaría un evento que ocurre en un punto particular del espacio. Bajo las suposiciones 1, 2 y 3, se puede demostrar que el número de eventos que ocurren en la región R tiene una distribución de Poisson con parámetro λa donde a es el área de R (λ se interpreta como la densidad del proceso). Se trata ahora de un **proceso espacial de Poisson**.

Ejemplos:

- 1- Suponga que aviones pequeños llegan a cierto aeropuerto según un proceso de Poisson, con tasa $\lambda = 8$ aviones por hora, de modo que el número de llegadas durante un período de t horas es una v.a. Poisson con parámetro $\lambda = 8t$.
 - a) ¿Cuál es la probabilidad de que exactamente 5 aviones pequeños lleguen durante un período de una hora? ¿Por lo menos 5?
 - b) ¿Cuál es la probabilidad de que por lo menos 20 aviones pequeños lleguen durante un período de 2 ½ hs? ¿De que a lo sumo 10 lleguen en ese período?

Solución:

- a) Sea la v.a. X : “número de aviones pequeños que llegan a cierto aeropuerto en una hora”
Entonces $X \sim P(8)$. Por lo tanto

$$P(X = 5) = e^{-8} \frac{8^5}{5!} = 0.0916$$

O también usando la tabla de distribución acumulada para la Poisson

$$P(X = 5) = P(X \leq 5) - P(X \leq 4) = 0.191 - 0.099 = 0.092$$

Y la probabilidad de que lleguen al menos 5 aviones será

$$P(X \geq 5) = 1 - P(X \leq 4) = 1 - 0.099 = 0.901$$

- b) Sea la v.a. X : “número de aviones pequeños que llegan a cierto aeropuerto en 2 ½ horas”

Entonces $X \sim P(8 \times 2.5)$ es decir ahora $\lambda t = 8 \times 2.5 = 20$

$$P(X \geq 20) = 1 - P(X \leq 19) = 1 - 0.470 = 0.53$$

↓
por tabla de F.d.a.

Y por último calculamos por tabla $P(X \leq 10) = 0.010$

2- Se supone que el número de defectos en los rollos de tela de cierta industria textil es una v.a. Poisson con tasa 0.1 defectos por metro cuadrado..

- a) ¿Cuál es la probabilidad de tener dos defectos en un metro cuadrado de tela?
- b) ¿Cuál es la probabilidad de tener un defecto en 10 metros cuadrados de tela?
- c) ¿Cuál es la probabilidad de que no halla defectos en 20 metros cuadrados de tela?
- d) supongamos que el número de defectos está relacionado con la máquina que produce la tela, debido a desperfectos de la máquina el número de defectos varía en ciertos tramos del rollo. ¿Se puede asumir que el número de defectos sigue una distribución de Poisson?

Solución:

a) Sea X : “número de defectos en un metro cuadrado”. Entonces $X \sim P(0.1)$ pues

$$\lambda a = \lambda \times 1 = 0.1 \times 1 = 0.1$$

$$P(X = 2) = e^{-0.1} \frac{0.1^2}{2!} = 0.004524$$

b) Si X : “número de defectos en 10 metros cuadrados”. Entonces $X \sim P(1)$ pues

$$\lambda a = \lambda \times 10 = 0.1 \times 10 = 1$$

$$P(X = 1) = e^{-1} \frac{1^1}{1!} = 0.3678794$$

c) X : “número de defectos en 20 metros cuadrados”. Entonces $X \sim P(2)$ pues

$$\lambda a = \lambda \times 20 = 0.1 \times 20 = 2$$

$$P(X = 0) = e^{-2} \frac{2^0}{0!} = 0.135$$

d) NO se puede asumir que el número de defectos sigue una distribución de Poisson, ya que las suposiciones que debe satisfacer un proceso de Poisson no se cumplirían.